**Endocrine Disruptor Methods Advisory Committee**
**First Plenary Meeting**
**April 26-28, 2005**

**Meeting Summary**

On April 26-28, 2005, the U.S. Environmental Protection Agency (EPA) convened the first meeting of the Endocrine Disruptor Methods Advisory Committee (EDMVAC). The meeting objectives included:

1. Review EDMVAC mission statement and operating procedures
2. Update on EDMVAC work plan
3. Review and discuss:
    a. Steroidogenesis (Tier 1 Assay)
    b. Uterotrophic (Tier 1 Assay, OECD)
    c. EPA Fish Screen Multi-Chemical Studies (Tier 1 Assay)
    d. OECD Fish Screen Phase 1B (Tier 1 Assay)
    e. Amphibian Metamorphosis Phase 1 Report and Phase 2 Draft Plan (Tier 1 Assay, OECD)

Copies of presentation slides and other materials distributed at the meeting may be obtained by contacting Jane Smith at smith.jane-scott@epa.gov or 202/564-8476. Many of the materials are also available on the EPA website at http://www.epa.gov/scipoly/oscpendo/. EPA established an administrative record for this meeting under docket control number OPPT-2005-0012. The official public docket is the collection of materials available for the public at the EPA Docket Center, Rm. B102-Reading Room, EPA West, 1301 Constitution Ave., N.W. Washington, DC. The EPA Docket Center is open from 8:30 a.m. to 4:30 p.m., Monday through Friday, excluding legal holidays. The EPA Docket Center Reading Room telephone number is (202) 566-1744 and the telephone number for the OPPT Docket, which is located in EPA Docket Center, is (202) 566-0280 or electronic access at http://www.epa.gov/edocket.

## I. Welcome and Opening Comments

Margaret Schneider, Deputy Assistant Administrator for the EPA, Office of Prevention, Pesticides, and Toxic Substances (OPPTS) welcomed the EDMVAC and members of the public for Susan B. Hazen, Acting Assistant Administrator for OPPTS. Ms. Schneider thanked members of the EDMVAC for their time and service to the Committee. She outlined three charges for the EDMVAC:

- Bring the best science to bear on the development of assays by the EPA's Endocrine Disruptor Screening Program (EDSP);
- Contribute scientific and technical expertise to EDSP's review of Tier 1 and Tier 2 science issues, protocols, and data; and
- Advise the EPA on the composition of a Tier 1 screening battery.

Ms. Schneider introduced Dr. Clifford Gabriel, Director for the EPA, Office of Science Coordination and Policy (OSCP) within OPPTS.

Dr. Gabriel expressed his appreciation to members for participating in the EDMVAC. He noted the enthusiasm within EPA and the EDSP about receiving input and advice from the EDMVAC. He explained that the EDSP is challenging both scientifically and programmatically. It is challenging scientifically because EDSP is attempting to develop assays to provide information specific to each chemical data with rapidly evolving science. Programmatically, it is challenging because of the number of offices and agencies involved, including the Office of Research and Development and OPPTS within EPA. Furthermore, there are other domestic and international organizations, all of which are simultaneously working through the complexity of validation.

## II. Introductions, Agenda Review, and Review of Previous Meeting Summary

EDMVAC members and key EPA's EDSP staff members introduced themselves. Dr. Gerald LeBlanc, EDMVAC Chair, welcomed EDMVAC members and explained that he and his Co-Chair would act as participating members of the Committee. He noted the role of the Co-Chairs is to facilitate information exchange from the committee to EPA. He asked members to call on the Co-Chairs when needed. Tom Osimitz, EDMVAC Co-Chair, commented that he is a new member of the EDMVAC (not on the previous EDMVS) and that he will therefore represent the new members of the EDMVAC. He asked that members ask clarifying questions when they need to, given the large volume and complexity of material.

Joseph Bailey, a Designated Federal Official (DFO) for the Science Advisory Panel (SAP) in OSCP, sat in for Jane Smith as the DFO at day one of the EDMVAC April 2005 meeting. Mr. Bailey explained that the DFO serves as a liaison between the Committee and EPA. The principal function of the DFO is to ensure that all rules of the Federal Advisory Committee Act (FACA) are met for the meeting. Mr. Bailey noted that this was a public meeting and that all background materials, presentations, and public comments from the meeting would be posted in the docket and on the website noted on page 1. He also mentioned that a meeting summary of the first EDMVAC meeting will be publicly available about six weeks after the meeting.

Maggie Rodriguez, Megatech Inc., provided information to EDMVAC members on the travel reimbursement procedures.

Juliana Birkhoff, senior mediator with RESOLVE and facilitator of the meeting, offered a brief introduction to her background and the role of RESOLVE in the EDMVAC process. Dr. Birkhoff reviewed the meeting agenda and logistics for the meeting. She emphasized that all members of the public must sign in and encouraged the public to sign up to reserve time to speak.

## III. Review of EDMVAC Basics

Gerald LeBlanc briefly reviewed basic information about the EDMVAC. Dr. LeBlanc reviewed a list of commonly used acronyms. He re-stated the charge of the EDMVAC, which is to provide scientific advice and guidance to EPA along the path to regulatory implementation of a screening and testing program for endocrine disrupting chemicals. He explained that consensus

EDMVAC recommendations to EPA would be ideal, but that consensus was not obligatory. Dr. LeBlanc described the process followed by EDSP toward the validation of assays: 1) detailed review papers, 2) protocol optimization and prevalidation, 3) inter-lab validation studies and 4) peer review.

Dr. LeBlanc also went over the assays EPA is currently developing. The Tier 1 screening assays under development are: amphibian metamorphosis; androgen receptor binding; estrogen receptor binding; steroidogenesis; aromatase; fish screen; Hershberger; uterotrophic; pubertal female; and pubertal male. There are also Tier 2 multi-generational assays under development. These assays are designed to assess the toxicity of endocrine-active chemicals over two generations of an organism by examining effects on the offspring of exposed individuals. Dr. LeBlanc concluded his presentation by reviewing the main players in the endocrine disruptor methods validation process. The players include EPA program administrators, EPA technical personnel, EDMVAC members, the RESOLVE facilitation team, the Organization of Economic Co-operation and Development (OECD), laboratory contractors, and the public.

*Discussion/Clarifications*

A member commented that the concept of prevalidation is important and that an integral part of prevalidation is a prediction model for how a chemical is going to react. An in vivo assay should develop what positive and negative results are setting the standard within prevalidation for the assay. The member said that a hypothesis about how the assay will act should be established in prevalidation so that validation is based on what was learned during prevalidation. This would create a clearer link between assay development and validation. Dr. LeBlanc noted that the member was suggesting establishing an expectation for the assay in prevalidation that should then be met in validation. Other members agreed that the purpose of prevalidation is unclear and often misunderstood. One member was concerned that prevalidation has become phase one of validation in the EPA process and validation, phase two. He stated that prevalidation should be for protocol optimization, testing transferability of the assay and developing a prediction model. Another member added that without a prediction model of how an assay is supposed to work, one cannot determine whether or not an assay worked properly at the validation stage. A member asked if prediction models were the same as acceptability criteria. Another member commented that they are similar, but prediction models are meant to set standards of what constitutes positive and negative responses in an animal.

A member pointed out that ICCVAM developed agreement on definitions for validation and criteria associated with each stage (see: http://iccvam.niehs.nih.gov/docs/docs.htm or the Fifth Plenary of the EDMVS at http://www.epa.gov/scipoly/oscpendo/ assayvalidation/ mtg_072302.htm.) He suggested that this document be provided to EDMVAC members. Members of the EDMVAC agreed that having these documents for reference would be helpful. An EPA staff member commented that EPA would like to devote much of the next EDMVAC meeting to discussing validation. Dr. LeBlanc suggested that these documents be collected and distributed to EDMVAC members so that they can think through validation issues to prepare for the next EDMVAC meeting. He noted that all members of the Committee seemed to agree that the EPA and EDMVAC must address validation process questions.

## IV. Review of EDMVAC Work Plan

Mr. Gary Timm, Dr. Les Touart and Mr. Jim Kariya of EPA's EDSP presented the EDMVAC Work Plan to Committee members. The Work Plan reviewed the assays under development by the EPA's EDSP, showed the progress of the EDSP, reviewed the status of each Tier 1 and Tier 2 assay and outlined projected next steps for the assays. The Tier 1 screening battery assays recommended for consideration by the original Endocrine Disruptor Screening and Testing Advisory Committee (EDSTAC) and under development by the EDSP include the following:

- Steroidogenesis
- Pubertal – female
- Fish Screen (aka Short Term Fish Reproduction)
- Amphibian Metamorphosis
- Estrogen/Androgen Receptor Binding
- Hershberger
- Uterotrophic

The Tier 2 assays recommended for consideration by EDSTAC include the following:

- Mammalian 2-generation
- Avian 2-generation Reproduction
- Amphibian Development and Reproduction
- Invertebrate (Mysid) Lifecycle (2-Generation)
- Fish Lifecycle (2-Generation)

The alternative screening assays recommended for consideration by EDSTAC include the following:

- Aromatase
- Pubertal – male
- Adult 14-day intact male

*Discussion/Clarifications*

Several members posed questions to EPA staff about the purpose of the methods validation process. One member asked about the distinction and relationship between screens and tests and how the EPA will use them. Mr. Kariya noted that EDSTAC did not intend for screening assays to determine adverse effects or dose response. The intent is that Tier 1 screens will provide clues as to what Tier 2 tests to run. A participant affirmed that this was the intent of the EDSTAC recommendations. Another participant reminded the Committee that EDSTAC was a previous advisory committee and published its recommendations, but the recommendations do not constitute a formal U.S. government policy.

Mr. Timm commented that the EDSP has spent its time developing assays, identifying chemicals to test, and working on a regulatory framework. Mr. Timm stated that EDSP has not yet precisely determined how all of the pieces fit together. Therefore, the logic of how to move from Tier 1 to Tier 2 is not clearly outlined and is an EPA policy issue to be decided. Mr. Timm asked that the EDMVAC focus on validation questions related to the assays under consideration at this EDMVAC meeting because policy questions must be addressed in other forums.

Several EDMVAC members commented that points raised about the purpose of validation would be important to address in the near future. Multiple members stated that they would like EPA to clarify the purpose of screens and definitive tests. One member also requested information about EPA's proposed method for developing a screening battery. Another member recalled from participating in the EDMVS that Tier 1 assays are simply supposed to give a yes or no answer about whether a chemical is endocrine-active. Members voiced that a recapitulation of the EDSTAC recommendations and a common working understanding of the purpose of screens, tests and validation would be useful so they could effectively advise the EPA.

Dr. Birkhoff noted that the discussion of the EDMVAC Work Plan seemed to generate short-term and long-term questions. In the short-term, EDMVAC members asked for clarification about EPA's definition of validation. Over the long-term, it sounded as if members were recommending that EPA develop a clear policy about how assays will be used to screen and test chemicals.

Dr. Clifford Gabriel, Director for EPA OSCP, summarized the points he heard EDMVAC members raise during discussion of the EDMVAC Work Plan:

1. How will the Tier 1 battery be used?
2. How can one confirm that a Tier 1 screen is working?
3. What chemicals will be used for testing?

Dr. Gabriel explained that Tier 1 is intended to only be a screen, meaning that in general, if results are positive then the chemical will advance to Tier 2 testing and that if it is negative, the EPA would decide what to do with that chemical. He noted that the assays should be considered independently of one another before they are considered as part of a battery. Dr. Gabriel stated that EPA will provide the list of the reference chemicals to be used in the EDSP validation at the next meeting.

Several members raised further concerns about the development of the screening battery and the chemicals that will be used to test it. A member inquired how the EPA was developing the screening battery and how they will assess its usefulness. He asked how many candidate assays there would be and whether all test chemicals will be run through all candidate assays. Dr. Gabriel responded that EPA development of the screening battery will be an iterative process. He stated that EPA will piece together the screening battery in the most effective way it can, based on the advice of the EDMVAC. This set of chemicals will be run through assays to help refine them. This aspect of assay assessment will take place outside of the actual validation process. He pointed out that the reference chemicals selected at this point were only chosen to test the performance of assays, not to test those chemicals. Mr. Timm added that if an assay completely failed than it would no longer be considered for the battery. He emphasized that EPA aims to develop an efficient battery comprised of a complementary set of assays and that redundancy might be needed to ensure scientific validity and reliability.

Several members offered comments about ways to optimize the Tier 1 screening battery. One member stated that it is critical to run chemicals that will challenge the performance of the assay so that confidence in its future performance is generated. Another member suggested that a common set of chemicals be used to test the performance of all assays so that divergence and/or complementarities will be evident. Another member noted that EPA must have a sense of

whether the assays are producing false negatives or false positives to properly assess the battery. A member reminded the Committee of the distinction between screens and tests. He reiterated that Tier 1 screens are only meant as a pass/fail to determine whether a chemical goes onto Tier 2. Therefore high numbers of false positives are not a problem in Tier 1 because the chemical will still be subjected to definitive testing in Tier 2.

A member asked EPA staff to explain how chemicals are currently being chosen for assay development. Mr. Timm acknowledged that choosing reference chemicals is a major challenge of validation and noted that finding definite negatives is particularly difficult. He explained that EPA has asked experts for nominations of chemicals that have demonstrated endocrine activity and for which there is quality data. Mr. Timm said EPA aims to choose chemicals to test across assays that include positives and some negatives. He noted that an important problem is that nobody has defined how many reference chemicals are adequate. In order for EPA to show the reference chemicals selected/proposed, a member suggested creating a matrix that shows what chemicals are being used and what endpoints are being examined. Mr. Kariya noted that EPA is developing such a matrix and will present it in the future.

## V. Update on Steroidogenesis

Mr. Timm introduced the update and discussion of the steroidogenesis assay. He asked the Committee to consider if EPA should proceed with validation of this assay, or are additional studies necessary before validation can begin. If more studies are needed, what does the Committee recommend? The assay is in the prevalidation stage, but not quite ready to move into validation. EPA could come back to the Committee in the fall with a validation proposal, depending on the feedback received and ability to resolve any outstanding questions about the assay before then.

Dr. Jerome Goldman, Reproductive Toxicology Division/NHEERL/ORD/EPA, updated the Committee on the Steroidogenesis sliced testis protocol. He outlined considerations for selecting a screen to evaluate toxicant effects on steroidogenesis and diagramed the sliced testis protocol for in vitro sliced testis incubation. EPA evaluated the protocol through both a multichemical study in the lead laboratory and positive control study in multiple laboratories. For the first phase, EPA tested 10 compounds in the lead laboratory. Dr. Goldman described the effect observed and possible explanations for each compound. In phase 2, EPA selected five labs based on several criteria related to knowledge and experience. Dr. Goldman summarized the results of the multi-lab comparisons. EPA labs also looked at cytotoxicity in the assay, particularly in determining the ability to distinguish it from effects on steroidogenesis. In conclusion, Dr. Goldman reviewed the major strengths and weaknesses of the steroidogenesis sliced testes assay and performance criteria. He noted that the assay meets many of the screen selection criteria, except for the need for lower concentrations of exposure and that only modest numbers of samples can be assayed concurrently. He also offered several ongoing and suggested adjustments to the protocol.

*Clarifications/Discussion*

Members raised several questions and comments about identifying and measuring cytoxicity in the assay. Members inquired whether EPA considers it critical for an assay to distinguish specific effects on steroidogenesis from non-specific effects and cytotoxicity. Several members

expressed that this is a critical feature in an assay. Members added that the endpoint chosen for the cytotoxicity evaluations did not seem to work. Dr. Goldman noted that LDH is probably not the best measure and EPA may have to revisit that endpoint. The current protocol measures of the effect of cytotoxicity generally, not specific cytotoxicity to Leydig cells that produce testosterone. Leydig cells only comprise about 2% of all cells in the testes; therefore, Leydig cell specific toxicity could be missed by the current cytotoxicity assay resulting in false positives. The assay could result in false positives if non-specific effects on cytotoxicity are seen, which would be included in a weight of evidence approach. He added that compounds are available that could prevent cytotoxicity by closing pores in the membrane. A member noted that this may work for cells, but may not work for testes fragments. An EDSP staff member noted that EPA looks at this question by asking whether the chemical would be treated differently if it was found to be cytotoxic rather than an inhibitor of steroidogenisis.

One member recommended that the H295R cells be used instead to overcome the limitations of the sliced testes assays regarding cytotoxicity. Dr. Goldman explained that EPA has initiated studies with an experienced lead lab and three other participating labs for the H295R assay (although the data are not yet available). The H295R cell line is derived from a human adrenal cortical tumor. The cell line expresses steroidogenic enzymes and, at sufficient levels, provides detectable interruptions in the pathway; therefore, it may be useful to assess steroidogenesis inhibition.

A member suggested that to address cytoxicity problems, EPA should run a specific assay for C4 using specific substrates that could include different measures for mitochondrial functions for cytoxicity. Another member suggested staining specifically for Leydig cells within the testes.

On the topic of statistics, one member commented that the assay is not very sensitive given the high coefficient of variation (CV) and asked whether EPA conducted a power analysis. Dr. Goldman responded that EPA conducted a power analysis which showed a power of 80 and above with an N of 6.

Other member comments and suggestions included changing the title of assay to accurately reflect the use of testis fragments rather than sliced testis and to use both testes. Several members noted that the presentation and array of data for the assay was useful.

Dr. Goldman offered the following clarifications on other specific technical questions:

- Regarding the temperature used in the protocol, evidence from the lead lab showed 34ºC as optimal because that is the normal testicular temperature, but no real differences are observed between 36ºC and 34ºC degrees.
- The assay included verapamil because the EDMVS suggested that compound.
- Dose levels are based on what is suggested in the literature and what was used in assays previously. Differences from the Laskey data could come from different dose ranges.
- No dose spiking was done in the assay.
- Using both testes may introduce more variability.
- No data on H295R cells is available to share at this time, although a lead lab and three other labs are participating in the work.
- Some chemicals were selected because previous work showed impact on testosterone, and some were selected because of EPA's experience with the chemical.

EPA suggested the following steps to refine the assay:

- refine the protocol to reduce variability
- run chemicals at lower concentrations
- look to other methods for measuring cytotoxicity
- investigate using pore-blocking agents to inhibit apoptotic cascade.
- Run 3-5 weak, moderate, and strong positive chemicals at five dose levels to obtain clearer evidence on the sensitivity of the assay.
- Establish an upper limit dose that is relevant to the biology of the organism.

The Committee discussed the specific question of whether EPA should continue with the sliced testis assay. Members noted that cytoxicity is a major concern that would need to be addressed if the assay is pursued. Some members noted that it is premature to remove the assay from consideration. Others suggested that EPA hold on the assay, wait to get results from the H295R assay, compare the two assays, and discuss them at a future meeting in 2005. Some members noted that the foundation of the assay is flawed and should not go forward. One member added that the purpose of the assay is to develop a prediction model to make decisions based on the assay results.

At the conclusion of the discussion of the steroidogenesis assay, the Committee agreed to the following consensus recommendation to EPA[1]:

***The EDMVAC recommends that EPA not move forward with steps to refine the steroidogenesis sliced testis assay. The Agency should instead invest resources in exploring better methods for testing for cytotoxicity. EPA should also invest more resources in pursing the H295R assay, and in addition, explore other assay options if the H295R assay does not meet the Agency's needs.***

## VI. Closing Remarks – Day 1

Dr. LeBlanc summarized the highlights of day one. He pointed out that reaching consensus is a welcome outcome, but that it is not mandatory for EDMVAC to reach consensus on its recommendations. Dr. LeBlanc reviewed the main discussion points of the day and where the Committee and/or EPA stand on each:

- *Questions about prevalidation*
  - ICCVAM documents will be collected and distributed to EDMVAC members to inform discussion about prevalidation.
- *Questions about how Tier 1 assays will be used in decision making*
  - EPA expects that Tier 1 should provide a yes/no answer about moving a chemical to Tier 2 testing; decisions relative to test requirements and risk assessments will

[2] Consensus is defined in the EDMVAC operating procedures as, "… there is no dissent by any member of the EDMVAC. Another way of stating this is that 'all members at the table can live with the decision.'"

be based on weight of evidence.
    - EPA suggested that the EDMVAC consider each assay independently at this time, not how each will fit into a battery.
- *Chemical Concentrations used in validation of the assays*
    - EDMVAC members noted that if concentrations are too high, than tests are not biologically relevant; and
    - EDMVAC members advised that there be consistency in the types of chemicals used as well as the concentrations used.
- *Consensus recommendation on Steroidogenesis assay*
    - EDMVAC members reached a consensus recommendation that EPA put development of the steroidogenesis assay on hold because of its inability to clearly differentiate effects on steroidogenesis versus cytotoxicity.

## VII. Opening Remarks – Day 2

*Documenting EDMVAC Consensus*
Dr. Birkhoff reviewed a proposal for how the EDMVAC would document consensus agreements among the members, such as the recommendation on the steroidogenesis assay. The recommendation would be recorded in the meeting summary. The EDMVAC chair would also create a separate document such as a letter to the EPA Administrator documenting the agreements. Other federal advisory committees have used this approach to raise awareness of consensus agreements in a more formal method and to have products to show from committee discussions.

Dr. LeBlanc explained that the product documenting the steroidogensis recommendation from this meeting would be a three paragraph letter. The three-paragraph letter would include a description of the issue and background, a review of the committee discussion and concerns, and a consensus statement.

Dr. Birkhoff reviewed the suggested process for sending the letter. The chairs would draft the letter, working with EDSP and RESOLVE on content and format. RESOLVE will send the letter to the committee by email for review. If changes to the letter are needed, the committee will work out a process to discuss the changes. When members agree to the letter, the chair will send it to the Administrator. Dr. Birkhoff explained that the EDMVAC operating procedures define consensus as "can live with."

A member suggested an alternative method where members vote on the letter to indicate concurrence or non-concurrence. Another member suggested that consensus could be a majority opinion. Dr. Birkhoff clarified that the EDMVAC operates by consensus wherever possible as its goal. Dr. Birkhoff added that the operation procedures allow for the Committee to say members were near agreement and include alternative views. Dr. LeBlanc suggested that the Committee try to reach consensus with this first letter, then discuss needed refinements at future meetings.

*EDMVAC Operating Procedures*
Dr. Birkhoff reviewed the Committee's draft Operating Procedures. The Procedures include provisions for the EDMVAC purpose, objective and scope, composition, alternates policy,

meeting procedures, roles, agenda procedures, minutes and records subcommittee and workgroup procedures, meeting attendance, public comment, and decision making/agreements, and expenses and reimbursements.

Members did not offer comments on the Operating Procedures at this meeting, but were invited to comment before or at the next EDMVAC meeting.

## VIII. Background and Update on Uterotrophic Assay

*Introduction by Mr. Gary Timm*
Mr. Timm provided background on the uterotrophic assay. EDSTAC recommended the assay, which was later selected by the OECD's Endocrine Disruptor Testing and Assessment working group as a priority for validation. He noted that it is an advantage to harmonize the assay internationally with the OECD work. The OECD Peer Review Panel's draft report on the assay contains a wide spectrum of views. The panel did not reach consensus that the assay was validated. EPA has not provided the report to the Committee because it has not yet been released to the public. EPA will come back to the Committee when the peer review report is available. A member commented that the peer review process for this assay did not contain the usual opportunity for public input. EPA staff responded that the Committee will also receive a description of the peer review process.

*Presentation by Dr. Willie Owens*
Dr. Willie Owens, Procter & Gamble, provided an overview of the validation of the uterotrophic bioassay. He described the general history of the assay from the 1930s to the recent timeline and status. He summarized the validation process, including purpose, assay and biological purpose, protocol, chemical selection, and reproducibility. Dr. Owens also described the statistical approach, prediction model and performance of the assay. He addressed rat diet issues with the assay, which have raised concerns in the past regarding the presence of phytoestrogens. He noted limitations of the protocol and identification of problematic data. In conclusion, Dr. Owens stated that the uterotrophic protocols were adequate, the reproducibility was excellent, and the predictive capacity is very good. Chemical selection was appropriate, but the need for other negatives should be discussed. He also concluded that the statistical approach was appropriate and functions well and the number of labs and animals used was reasonable.

*Clarifications/Discussion*
Overall assay performance and protocol
Several members commented that the data demonstrate that the assay works and identifies estrogen agonist activity. If there is any intent to use the assay to identify anti-estrogens, one commenter suggested that additional work is necessary. Other comments on the limitations of the assay recommended the inclusion of a weak positive control to show that the assay is robust enough to detect a weak signal. Other members suggested that a standard protocol to generate data is needed, which should be based on standardized conditions such as diet.

Tests should not be so specific that no lab can replicate another; some flexibility is appropriate for a biological assay of this type, but not for some in vitro tests.

One member commented that EPA should not be looking at this assay as an anti-estrogen assay because only one anti-estrogen was used. Dr. Owens clarified that a technical issues arises with the anti-estrogen, specifically partial agonism. One supplier did demonstrate a pure anti-estrogen, which was used in the assay. He also explained that data on the one anti-estrogen used are in the Phase 1 report.

In expressing general support for the assay as a screening mechanism, a member referenced data from Ashby's study that DHT has an uterotrophic effect and that the effect is blocked by flutamide. The member added that the exact mechanism is a secondary question to be addressed during Tier 2 or further testing.

Animal welfare

Regarding animal welfare issues, several members noted that the choice of protocol has implications from an animal welfare perspective and that the assay should use the fewest number of animals necessary to achieve an effective assay. Several members suggested that the assay use the intact animal; that protocol shows similar results as others and avoid subjecting animals to ovariectomy (OVX). Another member raised concerns that some labs used animal bedding and others did not, which could lead to differences in stress levels and hormonal effects. He suggested that guidance is needed to control for this variation. Members also inquired about the use of immature animals. Dr. Owens responded that in international meetings, differences existed in the preference for using immature intact animals or OVX animals. If the assay is a formalized test, treaty obligations would exist for the mutual acceptance of data. In that case, if OVX was used with a compound, EPA would likely be formally bound to accept the OVX data. One member noted that more discussion is needed on data to support conclusions that no difference exists in the route of exposure for immature and OVX animals.

In response to questions about the age of the animals, Dr. Owens clarified that the animal supplier provided the labs with animal age information. No consistent problems were seen with animal age, although some animals entering early puberty will always be a problem. The animals were weaned at 17 days.

Dose setting

Members raised several questions about dose setting in the protocol. In response, Dr. Owens explained that labs were given specific doses so comparisons could be made. The doses used in the assay were within standard toxicological practice and the protocol calls for three dose levels: moderate, highest effect, and no effect, in addition to the control. Members also suggested that guidance on dose setting could affect variation. Specifically, when testing an unknown chemical, labs need specific procedures for selecting maximum doses, especially for weakly-acting substances. Dr. Owens added that range findings were not done in many cases. The maximum tolerated dose was derived from the literature.

Negative controls

Some members raised concerns about the number of negative reference chemicals in the assay. One member commented that the assay performed pretty well overall with one negative chemical. Another member noted that with the addition of negative test materials, the majority of compounds that go through the screen will not have estrogenic activity. The member commented that the assay demonstrates an ability to detect weakly acting estrogenic compounds

with very high reliability. A weak positive control should also be included in the assay to demonstrate performance and portability.

Dr. Owens explained that in designing the assay, a compromise was reached to use one negative chemical to challenge the assay, dibutylphthalate. Dibutylphthalate was positive or statistically significant twice in Protocol B (one of which could be attributed to a loss in body weight) and once in Protocol C.

A member asked whether androgen, subject to aromatization, signaled false positives. Dr. Owens had noted that androgens are positive in the literature in the immature animals; they are in fact used for a direct aromatase inhibition assay by the pharmaceutical industry published over the last 15 years. With the ovaries present, the animals have sufficient aromatase activity to convert aromatizable androgens to active estrogens which elicit uterine growth. The absence of ovaries is an advantage for the protocols using ovarectomized animals. Dr. Owens responded that the uterotrophic is not perfectly specific for estrogens. If you use a steroidal aromatizable androgen, you will get a response and, hopefully, steroidal aromatizable androgens would not be required by regulators to be screened in the uterotrophic assay.

One member commented that the need for looking at toxicogenomics is unclear. Dr. Owens responded that toxicogenomics can be used to help determine whether labs can predict or obtain the same or similar dose response from an in vitro assay as other assays.

Phytoestrogens and diet

Several members noted that animal diet is an important feature of the assay protocol. On the issue of phytoestrogen in diets, members noted that evidence for low-phytoestrogen diets is well documented (such as the Thigpen study). Levels in animal feed can vary highly even from the same manufacturer. Phytoestrogens can have profound effects, such as advancing timing of the vaginal opening and possible effects on the uterus.

Other members added that diet needs to be considered and diet optimization would be helpful, but are probably too difficult considering the variability of the assay. In addition, if the protocol calls for controlling the diet of a juvenile animal, breeders also need to control the diets of the previous generation.

Dr. Owens explained that each lab preserved a sample of each lot of feed used during the study and that all lots were analyzed for phytoestrogens. The samples were sent to a (non-participating) lab and analyzed for three phytoestrogens. He agreed with the comments of some members that diet producers need a common, verified method to analyze and track phytoestrogens.

Another member commented that dose selection resulted in sensitivity from low body weight gain, which led to false positives. An ANOVA analysis should be conducted where body weights and animals that come into puberty early should be taken out as outliers

Statistics and predictivity values

A member commented that the assay has remarkable robustness. Another member noted that the data from the assay shows good correlation, but kinetics also should be considered.

A member commented that while using fewer animals is good, six animals are too few to get valid statistics, and replicates should be used. Another member commented that the performance statistics are not strong, particularly in positive and negative predictability. A member also raised concern about the predictivity tables. The member stated that the statistics are flawed because the assay included only one negative compound, which biases the statistics and could lead to false positives and false data.

*Discussion of EDMVAC Input on Uterotropic Assay*
Members then discussed the process questions of whether the Committee should provide opinions on the assay validation now or wait until the peer review is available. Mr. Timm reminded the Committee that the EDTA unanimously held that the assay is reliable and relevant for the purposes of an in vivo estrogen screen and that the assay would not be used to identify anti-estrogenicity in the EDSP.

Those in favor of waiting until the EDMVAC can review the final peer review noted that the review will provide background for members who aren't experts in this particular field and that the Committee should see the review before making a final decision. Minutes of EDTA or other records could be provided to members to provide some information before the final report is available. One member suggested that if the Committee waits for the peer review, members should provide EPA with a list of critical questions for the Committee to address at its next meeting, such as negative chemicals, anti-estrogens, and the types of compounds tested in other assays. Another member suggested that the Committee can provide technical comments before seeing the peer review report. Mr. Timm noted that the EDTA response to the peer review will also be available. A member noted that the ECVAM Scientific Advisory Committee (ESAC), which advises European Committee for the Validation of Alternative Methods (ECVAM) may also conduct a peer review of the assay.

One of the concerns of the peer review panel was that the assay was positive for aromatized androgens.

Some members raised questions about how EPA will use the information from OECD and the relationship of the assay to OECD test guidelines. Dr. Owens reminded the Committee that the OECD is a collection of regulatory agencies and does not set policies. Mr. Timm added that this assay is one of several candidate assays for Tier 1. Differences among countries exist on how assays should be used. EPA's decision on an in vivo screen for estrogenicity will depend on assay performance in relation to the performance of other assays. EPA will use a weight of evidence approach in selecting the overall test battery.

Dr. Birkhoff explained that the EDMVAC could hold a conference call or a meeting in July to discuss this assay further if the peer review report was published by that time.

## IX. OECD Fish Screen Phase 1B and EPA Fish Screen Multi-Chemical Studies (Tier 1 Assay) and Associated Histopathology

*Introduction by Dr. Les Touart*

Dr. Touart provided a brief introduction to the presentations on the fish screen Phase 1B and

EPA Fish Screen Multi-Chemical studies. He explained that Dr. Gary Ankley would cover the background of the fish assay, work done to date, and its current status. The most recent work conducted on the assay was by a Battelle contract for looking at multi-chemicals with the fathead minnow and OECD Phase 1B work. Dr. Touart mentioned that there was ongoing work occurring as a continuation of Phase 1B activity due to some questions that arose, particularly in relation to histopathology. He noted that such questions were part of the reason for asking Dr. Jeffrey Wolf to present information on the histopathology endpoint to the EDMVAC.

Dr. Touart stated that the fish short-term reproduction assay is intended to be used by the EDSP as a Tier 1 screening assay to capture estrogen and androgen active substances. Information obtained from this assay will be utilized as a presumptive trigger for Tier 2 or definitive testing to collect more information. He emphasized that Tier 1 is not meant to define modes of action.

Dr. Touart framed the key questions for EDMVAC members to consider during the presentations on the fish screen assay. The questions were:

1) For the fathead minnow, are the available data adequate to demonstrate that the assay is capable of responding to estrogen and androgen active substances? If not, what more is recommended?
2) Primary endpoints considered for inclusion consist of vitellogenin, secondary sex characteristics, gonad histopathology, and fecundity. Should additional endpoints be considered for purposes of screening potential endocrine disrupting substances?
3) Given that the fathead minnow is a species of interest and fecundity is an endpoint of interest, what would constitute a "false positive" in terms of a response observed in the assay and warranting Tier 2 definitive testing? How should "false negatives" be addressed?
4) What additional data are recommended to demonstrate the validity of the screening assay for capturing potential adverse effects and triggering Tier 2 testing?

*Presentation by Dr. Gerald Ankley*

Dr. Gerald Ankley, US EPA, Office of Research and Development (ORD), National Health and Environmental Effects Laboratory (NHEERL), presented an overview of EPA's development of a fish assay for detecting endocrine disrupting chemicals. (As indicated above, copies of slides from Dr. Ankley's presentation may be obtained from the docket or EPA website noted on page 1.) Dr. Ankley first reminded members of the set of assays EDSTAC originally proposed for a Tier 1 screening battery in 1998. One of the assays recommended by EDSTAC was a fish gonadal recrudescence assay to test for estrogens and androgens. He described the test species chosen for the development of a fish assay: the fathead minnow. Dr. Ankley reviewed the basic procedure and endpoints for the proposed gonadal recrudescence assay and explained the drawbacks EPA identified with this assay which include high variability among individual in recrudescence (gonads changing from resting state to spawning state), lack of details on reproductive bioassay of fathead minnow , logistical complexity in manipulating recrudescence. Due to difficulties encountered with the proposed gonadal recrudescence assay, EPA chose to begin work on an alternative fish assay.

Dr. Ankley then detailed the short-term reproduction test chosen as an alternative to the gonadal recrudescence assay. The short-term reproduction test examines a similar suite of endpoints as recommended for the recrudescence assay – secondary sex characteristics, ovary/testis

development (GSI, histology), sex steroids and vitellogenin. Dr. Ankley reviewed the following aspects of the short-term method for assessing the reproductive and developmental toxicity of endocrine-disrupting chemicals using the fathead minnow:

- test context;
- species background and biology;
- test system(s) and exposure methods;
- conducting of the assay and evaluation of "apical" endpoints;
- methods for "diagnostic" endpoints;
- statistical methods; and
- test interpretation.

Dr. Ankley reviewed data generated by EPA with known endocrine disrupting chemicals (EDCs) including estrogens (E2, Methoxychlor), androgens (α, β-Trenbolone, Methyltestosterone*), anti-androgens (Flutamide, Vinclozolin), and steroid metabolism modulators (Fadrozole, Prochloraz*, Fenarimol*)[2].

Dr. Ankley then presented data from external evaluation of the 21-day fathead minnow short-term reproduction assay test designs conducted by Battelle. Battelle conducted one study to evaluate responses to four known EDCs using the 21-day fathead minnow protocol (Methoxychlor, 17β-Trenbolone, Flutamide, Fadrozole). Battelle also conducted a second study to evaluate responses to five suspect/weak EDCs using the 21-day fathead minnow protocol (Atrazine, Bisphenol A, p,p'-DDE, Perchlorate, Cadmium).

Dr. Ankley then briefly reviewed the role of OECD in Endocrine Disruptor Chemical (EDC) testing before explaining the external evaluation studies OECD performed on the short-term fish reproduction assay. The OECD Fish Drafting Group (FDG) considered a range of protocols for a fish screen assay including:

- Short-term (14-day) vitellogenin induction assay in juvenile fish;
- Moderate-term (45-80-day) assays focused on gonad development/differentiation;
- Short-term (21-day) non spawning assay; and
- Short-term spawning assay initiated with reproductively active adults with endpoints being secondary sex characteristics, gonad histology, and vitellogenin status.

Dr. Ankley outlined the test method and results for OECD evaluation studies on the Phase 1A non-spawning assay and the Phase 1B spawning assay. OECD studies used mature zebra fish (Germany labs), medaka (Japanese lab), and fathead minnows, as well. The overall OECD conclusions were as follows:

- General test design/endpoints are reasonable for detecting estrogens, androgens and inhibitors of steroid metabolism;
- Anti-androgens are more challenging to detect;
- Standardization guidance for histological interpretation is desirable; and
- Design may need species specific optimization for spawning activity.

Dr. Ankley explained that moving into the future; EPA intends to enhance the diagnostic capability of the assay via genomics. Another EPA goal with the fish short-term reproduction

[2] Asterisk indicates (*) chemicals with mixed MOA.

assay is to understand linkages across endpoints to adverse outcomes at individual and population levels.

*Clarifications*

Experimental design and assay protocol

A participant inquired whether EPA ran chemical tests during all the characterization studies for the fish short-term reproduction assay. Dr. Ankley responded that EPA did test water and chemical concentrations routinely during tests (every two weeks in the Duluth lab).

The fish assay was considered important to EDSTAC because of widespread occurrence of intersex fishes and conservation of endocrine process across vertebrate species. The fish assay also complements the information from the mammalian assay(s) and fill the range of taxonomic subjects and metabolic functions for the Tier 1 battery. Fish are likely to show differences from mammals in EAT activity. Also, fish receive more dermal exposure than other taxonomic groups (except amphibians).

There was a question about methyltestosterone producing a false positive. This is not necessarily a false positive; we simply don't fully understand the mechanism(s) of action of the chemical.

A member asked whether EPA had set limits in conducting their characterization studies as to what is classified as an outlier response. Dr. Ankley replied that limits had been informally set, such as not using tanks with individuals that did not reproduce during the acclimation phase. He stated that EPA has not emphasized used co-variance, but that this is an aspect of the assay worth looking at. The member noted that regression analysis can be used to set limits, but that it is not entirely satisfactory.

A member asked whether EPA has developed enough of a control database for the fish short-term reproduction assay, based on the pre-exposure period, to determine what the assay can and cannot detect.

A member commented that fathead minnows are normally held at 25 degrees C and 24-26 degrees C is ideal for the reproduction test with FHM rodents body temperature is normally at 34 degrees. The member asked whether temperature has been shown to have any effect on androgen receptor binding. Dr. Ankley noted that the test species have wide temperature ranges so they have not been studied at particular temperatures.

A member asked if there are any important qualitative or quantitative differences between binding in mammalian test species and the fathead minnow. Dr. Ankley responded that comparison studies across vertebrate species are within two to five-folds of one another. Based on this comparison it appears the receptor systems are similar between the two regarding binding. However, the data are not adjusted for the difference in temperature at which the two test species live.

A member asked if chemical concentrations in the test fish were measured. Dr. Ankley responded that chemical concentrations in the fish were measured in the EPA studies and that tissues were preserved from most studies.

A member asked what was done when a male fish dies in a tank during the course of the study. Is that tank of fish lost? Dr. Ankley explained that if there are two males to a group of females in a tank the result for antiandrogenic effects can be more ambiguous, but if one of the replicate male dies, you still have a male left for spawning. If both replicates die then the issue must be addressed. He noted that if females are lost, data can be normalized to account for that occurrence by looking at per spawning data as the basis for a fecundity estimate.

Data and statistics

A member inquired about how data generated from the Phase 1B evaluation of the 21-day spawning assay were being expressed. The member noted that the data for the study shows a 0.5 if spawning occurred in one tank and a 1 if spawning occurred in both tanks. The member asked if fecundity was expressed as a ratio. Dr. Ankley replied that he was not entirely familiar with the data manipulations done by Masanori Seki, the Japanese researcher summarizing the Phase 1B study. Referencing data on the test species medaka, the member noted that the fecundity percentages were useful. However, he stated that using percentages with two tanks and an N of two may cause problems for drawing conclusions of statistical significance. He said that his personal impression of the Phase 1B data is that there were clear endocrine-disrupting effects demonstrated by the assay, but that proper statistical analysis is needed. Dr. Ankley replied that discussions are ongoing about bringing in independent statisticians to help optimize the design of the 21-day short-term fish reproduction assay. He also noted that EPA recognizes the need to adjust sex ratios and increase replicate numbers to address the issues raised. Dr. Les Touart of EPA added that currently underway is a Phase 1B follow-on study for which statisticians were consulted. He explained that replicates were being increased to 4 and that holding conditions for the test species were being modified to address variability in spawning across labs.

Assay protocol

A member asked what is known about fertilization success rates and embryo viability of the fathead minnow. Dr. Ankley responded that both aspects are routinely examined in these experiments, fertility and hatch. He noted that no chemical has yet been found to adversely affect either. He stated that fertility is typically in the range of 94-97% and hatch is usually 85-90%. Dr. Ankley noted that although effects on these variables have not yet been detected, they are still worth considering.

A member asked whether fish have ever been grouped based upon days post-spawning to reduce variability. He suggested that because of natural variability in reproduction, females could be grouped according to when they reproduced (e.g., 1, 2, 3 days previous) as they come out of the experiment so that there are subsets within the data. The member recognized that this reduces the N per group, but that this could be a way to reduce natural variability in the data. Dr. Ankley replied that attempts have been made to develop methods like this, but that it was determined to be impractical because a group spawning design (4 females to a tank) was in use. This design makes it difficult to identify which female spawned and to decide when to take down a tank. Dr. Ankley noted that this problem could be addressed to a degree by using a paired spawning design, but that this option is also impractical from a testing perspective.

Endpoint sensitivity

A member inquired if the fathead minnow has demonstrated unique sensitivity in any endpoints used in the fish screen assay compared to the mammalian assay for detecting androgen or

estrogen type activities. Dr. Ankley responded that fathead minnows seem to be uniquely sensitive to two modes of action relative to mammals. In his opinion, the vitellogenin response to estrogens is unparalleled, even compared to the uterotrophic assay. He also said the fathead minnow seems to be uniquely sensitive to inhibitors of steroidogenesis such as aromatase inhibitors. Dr. Ankley pointed out that rats receive cyclical exposure to the test chemical while fish receive a constant delivery of the chemical over the course of the assay. Thus, comparing the sensitivity of rats and fish may not be useful because one cannot distinguish the source of increased sensitivity.

A member commented on Dr Ankley's statements about the unique sensitivity of fathead minnow endpoints. He pointed out that protein responses other than vitellogenin are not looked at in mice or rats. He suggested that if other protein responses were examined, it is possible that sensitivity comparable to the fathead minnow could be found. The member also noted that according to the fathead minnow data presented, sensitivity for vitellogenin is relatively low compared to some of the other endpoints such as egg spawning success. This indicates a range of variation in response to estrogens.

A member asked whether studies showed a direct relationship between changes in secondary sex characteristics and reproductive performance. He mentioned the possibility that there could be compensation for changes in secondary sex characteristics by increased pheromone production. Dr. Ankley replied that EPA is collaborating with a lab specializing in fish behavioral endocrinology at the University of Minnesota that is studying behavioral endpoints. This lab is trying to determine if there is compensation or identify the mechanism of action. Dr. Ankley explained that the lab is working with both estrogens and androgens and is also pursuing work with pheromones. He noted that it is possible that this work will identify endpoints to be used in an EDC assay.

A member observed that the fat pad weight relative to the body weight in the male fathead minnow had not been used as an androgenic endpoint. The member noted that the fat pad index has been shown in certain labs to be a more sensitive endpoint than the formation of tubercles. Dr. Ankley said he was familiar with the fat pad index and that a colleague had provided a protocol for this endpoint to EPA. He noted that EPA is developing a data set comparing fat pad and tubercle endpoint data. It appears that tubercle formation and the fat pad are somewhat correlated, but that the fat pad may be slightly more sensitive because it responds quickly. The fat pad also offers a quantitative measurement, i.e., weight. EPA is considering incorporating guidance into the fish short-term reproduction assay through OECD to add fatpad index and tubercle count as endpoints.

Anti-androgens

A member asked if the dose of flutamide had ever been increased to try and generate a clear response from an anti-androgenic chemical, or if that would approach toxic levels. Dr. Ankley explained that in the Battelle studies, the high doses were in the range of 500-600 ppb per billion. The Japanese had tested the medaka with flutamide but did not detect anything, so they bumped the dose up to about 1000 ppb. At flutamide concentrations higher than 1000 ppb, signs of toxicity become evident in the fathead minnow. A member commented that it is not necessarily a lack of endpoint sensitivity rather that more anti-androgenic chemicals need to be tested like linuron or hydroxy-flutamide. In general, there is a need for anti-androgenic assays.

It was noted that in mammalian systems, hydroxy-flutamide is the active anti-androgen at the level of the androgen receptor. Dr. Ankley responded that EPA has tested hydroxyl flutamide in vitro, but not in vivo.

*Presentation by Dr. Jeffrey Wolf*
Dr. Jeffrey Wolf, Experimental Pathology Laboratories, Inc. presented information on the histopathology results of the Phase 1B fish screen assay studies. (As indicated above, copies of slides from Dr. Wolf's presentation may be obtained from the docket or website noted on page 1.) Dr. Wolf detailed histopathology findings and described the tissue characteristics in males and females exposed to the following EDCs used in the Phase 1B studies: fadrozole; prochloraz; flutamide; 4tPP; and estadiol.

Dr. Wolf described the proceedings of the Second Meeting of the Fish Pathologists for the Validation of Gonadal Histopathology in the Fish Screening Assay - Phase 1B. The meeting took place at University of Heidelberg, Germany in November 2004. The goal of the meeting was to determine whether histopathology is a useful and feasible endpoint in a screening assay for EDC effects in fish. Participants aimed to address the following:

- Is histopathology a sensitive and discriminating endpoint?;
- Is histopathology an efficient, economical and reliable endpoint?; and
- Further refine use of histopathology as an endpoint (revise guidance document).

Dr. Wolf outlined the objectives of the Heidelberg meeting and reviewed progress made toward those objectives at the meeting. The two overarching objectives of the meeting were:

1) Evaluate similarities among laboratories and pathologists with respect to the histopathology results of the fish screen assay; and
2) Evaluate differences among laboratories and pathologists with respect to the histopathology results of the fish screen assay.

Dr. Wolf reviewed the consensus findings of Phase 1B studies by chemical and according to test species, endpoints, gender and across labs. He highlighted areas of consistency among laboratories and pathologists. He also drew attention to potential sources for inconsistency between labs and researchers.

Genuine differences in results across labs may have been due to:

- Exposure levels differed from nominal;
- Reproductively immature fish;
- Fish substantially older than stipulated; and
- High mortality in controls (cause undetermined).

Artificial differences in results across labs may have been due to:

- At least one major misdiagnosis;
- Extra-gonadal changes in medaka not universally observed;
- Previously unrecognized findings including different diagnostic terms used by pathologists, or diagnoses not made by all pathologists;
- Not all pathologists attended key 2003 Phase 1A meeting in Paris; and
- Guidance document was incomplete.

Dr. Wolf presented a set of suggestions to make the histopathology endpoint more accurate, efficient and cost effective. His conclusions stated that histopathology is a valid endpoint for EDC screening because there were clear exposure-related findings for four of five chemicals tested. Dr. Wolf closed with recommendations and other considerations for improving histopathology of the fish screening assay.

*Clarifications*

Histopathology definitions and specificity

A member asked if Dr. Wolf was discriminating between what happened in the oocyte and what happened in the follicle layer in his description of histopathology results for prochloraz-exposed medaka. He noted that a true physiological atresia does not occur distinct from what happens in the follicle cells. Dr. Wolf clarified what he meant by physiological oocyte atresia, which he acknowledged may not be technically correct. The member suggested doing a time course study to look at the steps of the atresia process over time. Dr. Wolf noted that part of his conclusions indicate pathologists use different terminology to interpret test results and that this causes inconsistencies in data.

A member inquired whether thickening seen in the follicle cell layers in the prochloraz-exposed medaka is strictly in the granulosa. Dr. Wolf replied that the thickening was primarily in the granulosa. The member followed up by asking whether there was also a change in chorion thickness. Dr. Wolf answered that such thickening had not been observed.

A couple members commented on the lack of specificity with histological endpoints such as atresia. A member noted that atresia can occur due to a variety of factors such as a decline in vitellogenin or if spawning stops and vitellogenin remains high in females. In other words, atresia is not necessarily caused by exposure to an estrogenic or anti-androgenic toxicant. Dr. Wolf agreed that increased oocyte atresia can be caused by factors other than EDCs, but stated that atresia is an important endpoint to include in an overall battery of diagnoses.

Assay protocol for histopathology endpoints

A member asked why it is quicker to score slides non-blinded. Dr. Wolf replied that it is faster because when scoring slides blinded, researchers tend to look for diagnoses that probably are not there because they have no reference. He stated that the fastest way to read a study is to look at control animals in comparison to high dose animals first, and read down from the high dose. The member asked whether it was part of the Good Laboratory Practices (GLP) to have another pathologist review histopathology slides also for quality assurance. Dr. Wolf responded that this step had not been proposed. The member suggested that this step be considered regardless of added cost.

Another member asked how the common problem of producing high quality slides could be overcome. Dr. Wolf explained that the guidance document detailed how slides should be read and how they should be created.

A member asked with which chemicals mineralization in the testis ova was observed. Dr. Wolf explained that mineralization was not chemical dependent and that it was a background finding. He stated that it may be related to how the fish are raised or the quality of the water they are kept in, but that it is not a significant factor. Another member inquired if there are differences in

sensitivity to androgens or estrogens in the development of testis ova among the species used in the short-term fish reproduction assay (fathead minnow, zebrafish, and medaka). Dr. Wolf replied that medaka is the most sensitive, then zebrafish and fathead minnows seem to be least sensitive. He noted that this may have been a factor because younger fish were used, whereas testis ova were not as much of a factor in the Phase 1B studies with reproductively spawning adult fish.

Practicality and guidance

A member asked whether it would be reasonable to re-train traditionally-trained veterinary pathologists to conduct the fish screen assay. He noted that these pathologists do not normally have much exposure to working with fish. Dr. Wolf conveyed that if the number of diagnoses were sufficiently narrowed that it would not take a great deal of training for a traditionally-trained veterinary pathologist to execute the assay.

One member recommended that all members read over the background materials on the fish screening assay and pay particular attention to the guidance document. The member praised the effort displayed in the guidance document and stated that the multi-lab process is necessary to advance the science. The member also commented that it will take a major effort to cultivate the necessary level of expertise across labs to properly conduct the fish screening assay.

*Discussion*

Dr. Les Touart began the discussion by reviewing key points from Dr. Ankley and Dr. Wolf's presentations. He highlighted that the short-term fish reproduction assay with the fathead minnow seems to be very promising. He emphasized that many of the problems associated with the OECD inter-lab process are related to the use of different fish species across labs or attempts to shorten or simplify the assay. Dr. Touart acknowledged that there were difficulties with the Phase 1A non-spawning assay. He also listed some difficulties encountered with the Phase 1B spawning assay:

- Exposure variations;
- Age variations;
- Spawning variations; and
- Issues with the control fish performance.

He noted that performance criteria are needed across all species. He mentioned some other points of needed improvement:

- Improve the spawning conditions of fathead minnow (phase 1B follow-on study);
- Conduct quantitative measurements on fecundity and qualitative;
- Modify and optimize protocol/guidance for histopathology;

Dr. Touart emphasized that EPA wants to determine what is most practical and feasible to include in the fish screening assay. He asked that the EDMVAC provide input to EPA on the following aspects of the assay:

- What endpoints to add, subtract or modify; and
- What more is needed to cope with false negatives and false positives.

Dr. Touart reminded EDMVAC members that EPA intends for the short-term fish reproduction assay to serve as a Tier 1 screen and to keep that in mind when thinking about what is practical

to include at that level.

Assay Endpoints and Purpose

Initially, several EDMVAC members voiced support for the suite of endpoints chosen (vitellogenin, fecundity, histology). Several members agreed that the fish screening assay is a reproductive test and that fecundity is therefore a key measurement.

As discussion continued, several members expressed concern about the purpose and breadth of the assay. Key concerns and questions raised by EDMVAC members during the discussion included the following:

- Is fecundity a worthwhile and necessary endpoint?
- Fecundity can be affected by factors other than EDCs and is therefore prone to false positives.
- What are the most sensitive, powerful endpoints?
- What level of confidence exists that the chosen endpoints, specifically histologic, are truly indicative of endocrine effects?
- If core mechanisms can be assessed with the fecundity measurement, then what value do vitellogenin, GSI, hormone, histopathology endpoints add?
- What is the actual purpose of the fish screening assay – estrogens and androgens or reproduction?
- Is the fish screen assay intended to be a yes/no screen for Tier 2 testing or an umbrella assay for information on a variety of endpoints?
- Should the fish screen assay be made a Tier 2 assay due to its breadth and practical requirements?
- How will the assay be used to trigger Tier 2 testing?

Dr. Ankley explained that fecundity measurements are valuable from a quality assurance and an animal use perspective. Fecundity is a measure of reproductive health and fish performance. He suggested that fecundity data could also be used to set up longer term tests if a Tier 1 screen were ever used as a definitive test. Dr. Wolf said it would be difficult to prove the specificity of histologic diagnostic endpoints until the protocols are run with negative controls.

A member commented that there are a variety of experiments EPA could run to test the range of histologic responses such as manipulating the diet of the test species. Multiple members said they would like to see more specificity in the histology of the assay. A member suggested that performance standards and guidance be developed for the histopathology endpoints so that both experienced and inexperienced pathologists will be able to identify effects accurately.

Several members suggested that EPA clearly define the purpose and narrow the focus of this fish assay. A member pointed out that the original vision was an assay for estrogen and androgen. The member suggested the assay is being converted from a reproductive screen to becoming an umbrella assay for other endocrine activities beyond estrogen and androgen. Another member was not convinced that histopathology provided any endpoints more sensitive than vitellogenin. Fecundity is a good quality assurance measure, but is not an endpoint to get a reproduction answer. Another member inquired why animals were used if the goal was to look at whether compounds induced vitellogenin because this could have been performed in vitro.

Dr. Les Touart of EPA's EDSP responded to these concerns and stated that the short-term reproduction study is meant to indicate adverse effects and would not produce enough information to do a full risk assessment in Tier 2. He explained that the fish screening assay is complex and comprehensive because it covers a number of mechanisms that would otherwise take multiple assays to address. However, it does not produce enough information to serve as a Tier 2 test. The question is what should EPA do with non-endocrine active compounds that nevertheless indicate potential adverse effects. In other words, are false positives relevant in their own right?

Dr. Touart went on to provide historical context on the short-term reproduction fish assay. The original EDSTAC recommendation called for a fish gonadal recrudescence assay, which covered the HPG axis as well as estrogens and androgens. Fish were used so that the assay would be comprehensive and cover several mechanisms, enzymes and activities for which no assays existed at the time. The fish gonadal recrudescence was abandoned because of problems associated with using the fathead minnow and conducting it in a small lab setting.

Dr. Touart added that reproductively-active fish produce eggs and fecundity is a noninvasive way to assess health and reproductive ability of fish. Fecundity will respond to endocrine active materials and also to non-endocrine active materials. Fecundity as a screening assay endpoint may be indicative of adverse effects and may provide decision criteria for EPA to determine whether Tier2 testing is necessary for a given compound. Dr. Touart suggested that if fecundity is affected, whether by endocrine-active compounds or non-endocrine active compounds, EPA would still want to send the compound to Tier 2 testing. He emphasized that this should not be considered mission creep because it is EPA's responsibility to be aware of compounds that cause adverse effects. If it is determined that the effect is not endocrine-related, then a decision will be made about how to move forward. Dr. Touart stated that because the assay is developed for the endocrine program, does not mean EPA can ignore other important information about fecundity effects. A member acknowledged that it is EPA's prerogative to use information about non-endocrine-active reproductive toxicants as it sees fit.

Members' comments about fecundity as an endpoint for the fish screening assay covered the following:

- Fecundity is measurable and a reasonable endpoint for purposes of quality control;
- If looking at whole animals, fecundity cannot be ignored;
- The assay should not be centered on the fecundity endpoint and great energy should not be spent determining how to measure fecundity for statistical significance;
- Fecundity can add strength to the assay;
- Fecundity can validate other assays that may be showing effects related to endocrine activity;
- Fecundity may show a stand alone effect which suggests the chemical is not an endocrine toxicant; and
- Reproductive data may be used to support Tier 2 conclusions.

Many members agreed that fecundity be included as an endpoint. Members noted a variety of factors that can affect fecundity, other than EDCs. A few members strongly disagreed with the inclusion of fecundity as an endpoint and requested a clear definition of how fecundity data will be interpreted if it is to be employed as a toxicity indicator.

Anti-Androgens

Some members recommended that hydroxyl-flutamide be run through the assay before flutamide is re-tested using all fish at a comparable age, all fish in spawning condition, and the fathead minnow in an appropriate group size. A member noted that a reasonable hypothesis is that fish do not metabolize flutamide to the active anti-androgen. Another member stated that testing hydroxy-flutamide could help determine whether it is the metabolite the fish are not producing or that there is no effect because it is not being metabolized to address flutamide. If hydroxy-flutamide does not work, there are other anti-androgens like linuron that could be tested. She suggested possibly running anti-androgens along with an androgen and see if there is suppression in the androgenically-induced response.

The member also mentioned another assay for anti-androgens that is done with another fish species called the stickleback. The species is common in Europe. Males produce a protein called spiggin that they use to build nests. EDCs have been demonstrated to affect production of spiggin so that females start to make it when exposed to androgens. The effect is blocked by anti-androgens. She noted that there is convincing data from this assay for vinclozolin, linuron, and phenethylithone.

Assay Protocol and Statistics

A member asked whether EPA measures the concentration of chemicals in the test species' target organs. Dr. Ankley replied that this is done most of the time, but not always in every lab. The practice is recommended in the guidance, but EPA cannot make it mandatory because it would raise the assay cost too much and some labs do not have the necessary instrumentation. The member suggested that tissue samples be preserved for future testing.

Several members expressed concern over the number of replicates employed in the assay design. Each tank serves as a statistical unit in the design. The labs used both paired designs and group spawning. One member pointed out that if the fish screen assay is used as a reproductive screen, then an N of 4 is not high enough to generate statistically significant findings. Another member cautioned against dividing tanks to get more replicates because that is pseudo replication, which can cause problems.

A member agreed with the notion that for steroid hormones the CVs were much lower for RIAs versus the kits.

## X. Closing Remarks – Day 2

Dr. Gerald LeBlanc, EDMVAC Chair, highlighted the main discussion points of day two. He noted that members discussed the meaning of consensus for the purposes of the EDMVAC and developed a model for documenting consensus recommendations to EPA. Dr. LeBlanc reviewed key points of agreement and concern among members about the assays covered on day two:

*Uterotrophic Assay*

- EDMVAC members agreed that uterotrophic is a robust assay with much work behind it, but that there are several issues that need to be addressed.
- EDMVAC members agreed they were not yet prepared to give EPA advice on the state of

the assay.

- Members agreed they need to examine the OECD peer review report before making recommendations on the status of the uterotrophic assay.
- EDMVAC members expressed specific concerns about the assay:
    - A variety of protocol standardization concerns such as questions about what compounds should be tested, at what dosages and setting criteria for these decisions; and
    - Concerns about the assay's ability to detect anti-estrogenic effects.

*Fish Screen Assay*

- EPA acknowledged aspects of the fish screen assay that need to be addressed such as spawning conditions, histopathology and practicality considerations.
- EPA acknowledged that clearer guidance is needed for the fish screen assay.
- Need to clearly define the use of the fecundity endpoint
- EDMVAC members noted that EPA should run more negatives through the assay to learn more about the susceptibility for false positives.
- EDMVAC members suggested that EPA amend the fish screen assay so that it can detect anti-androgenic activity.
- EDMVAC members discussed endpoints to add to the assay or modify:
    - Members agreed the vitellogenin endpoint is useful;
    - Members eventually agreed that the fecundity end point is useful and adds value to the assay because it is measurable and helps look at the health of the study animals; and
    - Members agreed that secondary sex characteristics are an important endpoint; and
    - A member recommended adding the fat pad index to the fathead minnow protocol to address anti-androgens.

## XI. Opening Remarks – Day 3

*Signing In and Timing of Next EDMVAC Meeting*

Dr. Birkhoff opened day three with a reminder for attending members of the public to sign in each day of a FACA meeting so that EPA can track attendance. Jane Smith, EDMVAC DFO, noted that the next EDMVAC meeting would likely be held the first week of August 2005. She emphasized that the timing was subject to change based on member availability, and the status of EPA research and necessary background materials for the meeting.

*Topic Teams and Meeting Materials Assessment*

Dr. Birkhoff led a discussion among EDMVAC members about managing meeting materials. Several members expressed satisfaction with the employment of topic teams to prepare for the first EDMVAC meeting. Each team was asked to focus on studying in-depth the materials for one topic on the agenda, while not spending as much time on the other topics. This reduces the level of work required of members to prepare for EDMVAC meetings, but ensures that all topics are thoroughly covered by at least some members so that they can help clarify points during plenary. One member noted that for the first EDMVAC meeting, the topic team approach worked well because of the large volume of background material, but that it may not be necessary for meetings involving less material. Ms. Smith noted that topic team assignments are

flexible and that members may trade assignments if need be. All members will be assigned to a topic team at some time during the course of EDMVAC work.

Members also provided other suggestions about how to better manage the large volume of materials distributed to them for plenary meetings:

- Large binders with all meeting materials in print are costly and not entirely necessary;
- Put meeting materials on CD – supply as much as possible in advance of the meeting and put the balance of materials on a second CD to be distributed at the meeting;
- Make some print versions of the meeting materials available;
- Make presentations available in print at the meeting and put all other background materials (scientific studies and data) on CD; and
- Give members an option of whether to receive materials in print or on CD..

An EDMVAC member also encouraged EPA to continue developing assay short stories. Ms. Smith noted that all the meeting materials are also available electronically via the web site and the docket noted on page 1.

Dr. Birkhoff reviewed the agenda for day three.

## XII. Amphibian Metamorphosis Phase 1 Report and Phase 2

*Introduction by Dr. Les Touart*

Dr. Les Touart, EPA's EDSP, briefly outlined the information to be presented on the amphibian (frog) metamorphosis assay and the questions EPA had for EDMVAC members. Dr. Touart noted that frogs are being employed by EPA as a general vertebrate model for evaluating thyroid active substances in the screening tier. He explained that much work within OECD and the national framework has preceded the current state of the amphibian metamorphosis assay. Dr. Touart asked EDMVAC members to provide advice to EPA about:

- Other endpoints EPA should include for a screening assay; and
- How to tie up any loose ends before moving the assay toward validation.

*Presentation by Dr. Sigmund Degitz*

Dr. Sigmund Degitz, EPA ORD, NHEERL, Mid-Continental Ecology Division, reviewed the history, rationale and current status of the amphibian metamorphosis-based thyroid assay. He covered OECD Phase I validation results, OECD Phase II validation studies and supporting studies. (As indicated above, copies of slides from Dr. Ankley's presentation may be obtained from the docket or EPA website.)

Dr. Degitz outlined the protocol for the original EDSTAC-proposed Tier 1 Frog Metamorphosis Assay. Analysis of the assay's performance showed that it was not ready to be used as a screening tool. Research needs for improving the assay were identified, including:

- Determine appropriate stage to maximize sensitivity yet minimize time on the test;
- Determine appropriate endpoints;
- Establish diagnostic specificity of the assay;
- Conduct studies with known Hypothalamus-Pituitary-Thyroid (HPT) pathway disrupters; and

- Investigate interaction of other endocrine pathways on HPT.

Dr. Degitz reviewed progress made on early research through June 2003 as U.S. and international labs worked on variation of the EDSTAC protocol. In June 2003, OECD convened the Amphibian Expert Group to plan inter-laboratory studies as an initial step in the validation process.

Dr. Degitz went on to describe the OECD Phase I validation process including the rationale, approach, participating labs, data generated for controls, and data generated with frogs exposed to PTU and T4. He also detailed conclusions about endpoint sensitivity from German, Japanese and U.S. labs.

Dr. Degitz then reviewed OECD Phase II validation studies. The objectives of these studies were to use a more standardized protocol, increase the number of participating laboratories, and increase the number of chemicals tested. He outlined the Phase II protocol and endpoints, and summarized the current status of Phase II studies. The statistical review was completed, histopathology guidance was in development, and tests were ongoing in several labs with results expected in fall 2005. Dr. Degitz concluded with a brief discussion of additional studies being conducted in support of the amphibian metamorphosis assay.

*Clarifications*

Data

Some members asked questions about the data Dr. Degitz presented during his presentation. One member asked for a brief overview of how statistics were done to look at significant differences across stages. Dr. Degitz explained that a Kruskal-Wallis test was performed, using a one way analysis of variance as a non-parametric test. If that test shows significance, then a Dunns multiple comparison was used. The control was compared to all the different treatments. Dr. Degtitz acknowledged that these are not powerful statistical tests. A member inquired whether the Kruskal-Wallis test assumes a similar distribution among groups. Dr. Degitz said it does not.

Noting that some information was missing from the data presented, a member inquired whether EPA anticipates receiving additional amphibian metamorphosis histology data from the Japanese lab. Dr. Touart responded that he did not expect to receive more data from Japan for Phase 1. He noted that Phase 2 will be initiated soon and that guidance for that phase will be more specific such that all labs are expected to provide full sets of data. There is also other follow-up work being done by Battelle with Fort Environmental Labs and a report will be generated soon.

Assay Protocol

A member inquired about basic environmental controls in place for the amphibian metamorphosis assay protocol. Dr. Degitz explained that all three labs consistently used a standard photoperiod and standard temperature on all tests in Phase 1 prevalidation. However, there was deviation between labs after that, for example, the Germans used artificial water and US used Lake Superior water. Although each lab used there own water, the revalidation data indicated the data were consistent across labs. Dr. Degitz suggested the data may align even more once feeding and exposure regiments are made consistent across labs.

A member asked how uptake occurs in the frog (and fish) assay, given the problem of some substances not being very water soluble. Dr. Degitz stated that uptake is mainly through the gills. The EPA lab in Duluth, MN avoids using solvents if possible and has developed various techniques to get substances into a solution.

A member commented that the frogs used in the both the 14-day and 21-day U.S. studies were much smaller than the other labs. Dr. Degitz clarified that this aspect of the data was an artifact. The Germans and Japanese were measuring from the tip of the nose to the tip of the tail while U.S. labs were measuring from the snout to the vent. Another member asked which measurement will be used in the future protocol. Dr. Les Touart replied that both the whole body and snout vent length will be measured at all labs.

A member asked whether a decision has been made about whether to run a 14-day or 21-day test with the intent of reducing the number of false negatives. Dr. Degitz answered that a final decision had not been made on this point within the OECD process, but he felt the 21-day assay (Stage 51) was more sensitive (examples, phenobarbital and PCN) and was agreed for the Phase 2 work.

A member advised EPA to be cautious about the breadth of the assay and to prioritize which endpoints are essential. He encouraged EPA to focus on endpoints that add value, are doable and that work reliably.

Histology

A member inquired whether there was homogeneity among the responses of frogs to thyroid gland stimulation. The member wanted to know if histology can be used confidently as weaker agonists and antagonists are tested. Citing recent data on phenobarbital as an example, Dr. Degitz reported that clear effects on histology were consistently observed across all animals that were evaluated. He noted that dose response had been conducted and that the lowest concentrations that affected thyroid gland histology in frogs, affected all the animals consistently. The member noted that rats are not as consistent and may show changes or virtually no changes from baseline follicular development.

A member highlighted that it is the gland that shows definitive thyroid toxicity. He recommended that EPA look at some other developmental toxicants such as a retinoid, to make sure there is no secondary effect on the gland. The member asked for an indication of the relative sensitivity of the T3 and T4 levels in the animals as compared to the glands. Dr. Degitz said he could not give an indication of relative sensitivity. He stated that EPA is working to decipher relative sensitivity, but that the Agency has encountered difficulties with kits like RIAs and ELISAs because of serum differences. He said that EPA is trying to reproduce a study with T3 and T4, originally published in a well-cited French paper. A member asked whether measurements are being done in serum or plasma with the RIAs and/or the ELISAs. Dr. Degitz replied that measurements are being done in serum.

A member asked whether thyroid hormone receptors were measured in the pre-metamorphic stage, before stage 54 or before stage 51. Dr. Degitz explained that a number of labs have conducted work on TR alpha and TR beta and their expression in *xenopus laevis*. What this work has shown is that prior to the onset of stage 54, there is primarily TR alpha situation and it

is believed that the rise in TR alpha catalyzes the onset of metamorphosis. Once the onset of metamorphosis occurs, there is a transition from TR alpha expression to TR beta expression. This is more of a direct transcription of the genes that get regulated. If the animal controls gene expression in the battery of genes, they get expressed by which receptor is present. Researchers believe that TR beta sets in around stage 51 or 52 and that is why the animals become responsive to T3 and T4.

*Discussion*

Discussion of the amphibian metamorphosis assay was initiated by Dr. Les Touart of EPA. He noted that U.S. labs participating in Phase 2 studies are adding a PBDE-type compound to the other three for testing. EPA chose to add this compound because it is also being used in the rat pubertal assay and because it was recommended that EPA look at developmental toxicants with the frog assay. Dr. Touart also reminded members that EPA is developing clear histopathology guidance to ensure the use of standard and optimum procedures for collecting morphological and histological samples and for evaluating pathology. He also re-stated the key advice EPA sought from EDMVAC members about the frog assay, given the planned OECD Phase 2 validation work:

- What additional data are recommended, if any, to demonstrate the validity of the screening assay for capturing presumptive thyroid active substances and triggering Tier 2 testing?
- What additional endpoints should EPA consider?
- What additional compounds should be run through the assay?

Assay Protocol

Several members agreed that the frog assay has significant potential and expressed support for its use. However, EDMVAC members made several suggestions related to the protocol and proceeding with the validation process. Member suggestions included the following:

- Gain more experience and characterize normal frog development (specifically of T3, T4, and histopathology) for comparison to the exposed frogs in order to detect variation;
- Optimize the protocol before standardizing it;
- Specify a standard range for starting weight and length of test individuals to reduce variation (see OECD Guideline 215);
- Develop a prediction model with clear criteria;
- Develop a prediction model for identifying effects on the thyroid at different developmental stages and test the model in inter-lab studies.
- Develop clear guidance for standardization of the histopathology of the thyroid;
- Assess some chemicals that are weak agonists and antagonists to be sure the assay will detect them;
- Include chemicals that induce the induction of UDP-glucuronosyltransferase activity by the two pathways: aryl hydrocarbon receptor (AHR) and the pregnane X receptor (PXR);

Concern was expressed that this one assay is being relied on by the Agency to predict thyroid effects in humans from an amphibian. There are differences in metabolism and route of exposure to chemicals, for example. Dr. Touart reminded members of the EDSTAC concept of the Tier I battery and that we are considering other mammalian assays to address the thyroid axis, i.e. the pubertals, and an alternate adult male assay. There may be some question as to the

robust nature of the thyroid endpoints in the pubertal assays in comparison with the amphibian assay to determine whether one is satisfactory for the other or whether we need both assays to act complimentary to each other. We (EPA) need to be aware of how predictive the amphibian assay may be for mammalian systems.

Dr. Les Touart conveyed that a key goal of the OECD Phase 2 frog assay studies is to standardize the protocol. He emphasized that refining protocols and developing performance criteria are key aspects of validation and of OECD Phase 2 studies. He reiterated that the Phase 1 results were quite similar across labs. With further standardization, it is presumed that Phase 2 data will demonstrate greater consistency. Dr. Touart noted that technical histopathology guidance for tissue collection and preservation has already been drafted. Guidance for reading the tissues is under development. EPA's Duluth lab has the lead for communicating information across labs on these points.

Dr. Touart asked for EDMVAC members to recommend thyroid agonists to be tested in the future. A member commented that actual antagonism or agonism of the receptor can affect thyroid function. A member suggested that chemicals that affect the thyroid by induction of UPDs and enhanced clearance of the hormone be included in assay validation. The member noted that induction of UPD is typically mediated by two different receptor-mediated pathways. He also noted that consideration should be given to chemicals that induce these enzymes by way of the AHR and PXR receptor.

A member advised EPA to keep in mind that kits for T4 analysis that are purchased or made in labs themselves will not meet the criteria for studies required to be done under GLP conditions.

Statistics

A member noted that although there are differences evident in the frog assay data, they do not show up as statistically significant. The member expressed concern that this may indicate that the assay is not sensitive. He recognized that a non-parametric test was employed, but suggested trying another statistical method also. Dr. Touart replied that OECD consulted with statisticians about Phase 2. One possible change to the protocol is to increase the number of replicates during the validation exercise. The final version of the assay may not need as many. Considerations for a modified statistical design include using a larger number of controls with higher concentration and fewer replicates. Dr. Touart explained, however, that this unequal design would require difficult changes to existing lab infrastructure so Phase 2 is proceeding with existing infrastructure.

A member stated a preference for applying the Dunn's test rather than using the Kruskal-Wallis test. He considered the Kruskal-Wallis test to have less value since we are interested in comparing treated animals against controls. He also suggested the Jonckheere Terpstra test in order to gain greater sensitivity in the statistical comparisons between the treatments and the control. He also suggested developing a flow chart or decision diagram to illustrate the rationale of chosen non-parametric statistical approaches. Dr. Touart replied that the Jonckheere Terpstra test was suggested in consultations with statisticians about optimizing the frog assay statistics. He mentioned that he would distribute copies of the draft statistics document that resulted from the consultation.

Tier I Battery Concerns

A member advised EPA to work on coordinating its battery of assays for EDSP with other international regulatory initiatives (Europe and Japan). He pointed out that coordination would reduce duplication of data for thyroid and/or other endpoints and would reduce use of animal resources. Dr. Touart explained that there is an ongoing effort within the OECD to review thyroid assays in general and that alternative approaches to thyroid are being compared and scrutinized. The U.S. produced a detailed review paper on thyroid assays also. Dr. Touart stated that EPA will adhere to the Mutual Acceptance of Data (MAD) agreement. He also said that EPA will make a decision about the optimal tool for thyroid with the international context in mind.

One broad concern expressed is regulators around the world are building differing [ED] test systems and if these data are not shared among all the regulators, large scale over-testing may result.

One member expressed the importance of the sensitive windows of development involved in a rapidly developing organism, i.e. the frog assay. This is the only assay involving a rapidly developing organism proposed as a Tier 1 screen and that fact is significant.

## XIII. Public Comment

At the conclusion of the deliberations on the second and third days of the meeting, members of the public were given the opportunity to provide comments. Slides of some of the individuals' comments may be obtained from the docket or website noted on page 1.

*Vincent Kramer, Dow AgroSciences*

Dr. Vincent Kramer provided comments on the fish screening assays on behalf of the American Chemistry Council (ACC) and Crop Life America. He stated that EPA should clearly define the purpose of the fish short-term reproduction test, that harmonization with OECD should be enhanced to achieve global concordance with future screening results, and that fish screening assays are estrogen/androgen screens. Specifically, in answering EPA's questions about the fish screen assay, he responded that the assay should be used to screen estrogen and androgen active chemicals. The screen as it is envisioned is not designed or intended for thyroid active substances. Dr. Kramer also noted that the apical endpoints in the assay may be confounded by toxicity mechanisms other than estrogen/androgen substances and increases the complexity of the assay. He added that the fish screen is too long to be a screen and recommended it be shortened to 7-14 days.

He concurs that the tiered screen approach could disaggregate the biomarker endpoints (for example, the vitellogenin, steroid, secondary sex characteristics, et cetera) in a first tier screen followed by apical endpoints (eg. fecundity, histopathology) in a second tier test. He pointed out that every case where there was an effect on fecundity there was an effect on vitellogenin as well as the reverse.

*Richard Becker, ACC*

Dr. Richard Becker provided ACC perspectives on the validation of EDSP assays. He

emphasized the importance of validation as well as the need for the weight of evidence approach to be clearly articulated and agreed upon. He listed several points explaining why validation is essential and how it should be used. Validation is required by the Food Quality Protection Act (FQPA) and the ICCVAM authorization Act of 2000. Validation establishes relevance and reliability and helps interpret results. Dr. Becker outlined the differences between screening, which provides a mechanistic understanding of the potential to interact with one or more components of the endocrine system, and definitive tests, which yield information on apical endpoints for use in risk assessment. He provided a diagram showing the relationship among methods, test systems, and prediction models. He also listed several points that validation should address and several common principles of method validation. He concluded by describing some validation design challenges, included reference agents, use of positive and negative agents, and the need for sufficient data. Overall, he stressed the need to create validation principles, build and use prediction models, and maintain transparency in the design process.

*Lisa Ortego- Bayer Crop Science*
Dr. Ortego provided comments on the amphibian metamorphosis assay on behalf of ACC and in collaboration with Crop Life America. She commented that the differences in the Phase 1 protocols in the assay make it difficult to compare endpoint variability and reproducibility between labs. She also stated that EPA needs to resolve questions on the length of the assay and stage of development. In addition, a non-thyroid active toxicant should be included in either Phase 1 or 2 to gauge specificity and relevance of the endpoints. On the question of whether the screen is redundant, she responded that EPA has other assays such as the pubertal and 15-day male assay that should all be compared. In order to make this comparison, we have to ensure the same chemicals are tested across these assays. PTU is a thyroid agonist that is available for comparison purposes, however, there does not appear to be a thyroid agonist. The sensitivity of the 3 assays should be compared. Lastly, add a non-thyroid active toxicant to the phase 2.

*Angelina Duggan, Crop Life America*
Dr. Angelina Duggan provided general comments on the direction of the EDSP assays and the EDMVAC. She commented that since the release of the EDSTAC report in 1998, science has advanced. She urged the Committee to consider the progress already made and focus on moving forward. The executive summary of the EDSTAC report provides a good overview of the strategy and plans agreed upon by that Committee. Dr. Duggan also acknowledged challenged in a lack of human health data. Finally, she recommended that the EDMVAC deal with the technical issues and commended EPA for its overall multi-stakeholder process for endocrine disruptors.

## XIV. EDMVAC Materials Management

Members briefly discussed the process of distributing materials to the Committee. Member suggestions for handling the volume of printed material including posting to a password-protected website accessible by members and having a complete CD of materials available at the meetings rather than a notebook. A notebook should be available for a few members who prefer hard copies to keep a record of the meetings.

Members supported the assay stories approach, noting that the stories are a useful way to track the history and progress of the assays.

Ms. Smith summarized a plan for future materials management. Members will receive CDs 2-4 weeks before the meeting with substantive background information. Thin notebooks with handouts and presentations and other last-minute materials will be provided to members at the meetings. A few notebooks of all the materials will be available at the meeting for reference and can be taken home by members who want hard copies.

## XV. Closing Remarks and Summary of EDMVAC April Discussions/Feedback

Dr. Birkhoff began the closing remarks by reviewing the agreements and recommendations reached by the EDMVAC during their first meeting. She highlighted that the Committee did reach a consensus recommendation on the *sliced testis steroidogenesis assay*. The EDMVAC agreed that EPA should not move forward with the steroidogenesis assay and should pursue the H295R assay instead. EDMVAC members believe the cell-based H295R assay, which is also under development as a potential screening assay, possesses significantly greater promise than the sliced testis steroidogenesis assay. Dr. Birkhoff noted that this recommendation will be captured in the EDMVAC meeting summary as well as in a consensus letter to the EPA Administrator. The consensus letter will be drafted by the EDMVAC Chair and Co-Chair and circulated to members for review and comment. Once all members agree on the language of the letter, it will be submitted to the EPA Administrator.

Dr. Birkhoff cited that the EDMVAC did not make a recommendation to EPA on the *uterotrophic assay*, instead choosing to reserve judgment until the OECD peer review report is available to members for examination. She noted that EDMVAC suggestions and recommendations about the *fish screen assay* and *amphibian metamorphosis assay* will be captured in the meeting summary. Dr. Birkhoff pointed out to EDMVAC members that as EPA moves these two assays closer to validation, the Agency will want clearer advice from the Committee.

Dr. Gerald LeBlanc, EDMVAC Chair, expressed his appreciation to Committee members for a productive and successful meeting. He thanked EPA for providing valuable background stories on the assays under development and encouraged EPA to continue updating and improving the background stories. He concluded with a brief review of the main points discussed in relation to the amphibian metamorphosis assay on day three:

- EDMVAC members agreed that the amphibian metamorphosis assay is robust and that EPA should move forward with its development.
- EDMVAC members offered some suggestions about things to consider as EPA moves the assay forward:
    - Optimize protocol;
    - Consider developing performance criteria;
    - Improve standardization of protocol;
    - Improve statistical design;
    - Develop a prediction model;
    - Carefully choose reference chemicals and consider weak acting chemicals (agonists and antagonists);
    - Work to harmonize the amphibian metamorphosis assay with international efforts to detect thyroid effects;

- Acknowledge the unique value of amphibians in assessing toxicity due to their rapid development.

Tom Osimitz, EDMVAC Co-Chair, thanked EPA and RESOLVE for effectively bringing new EDMVAC members up to speed on the meeting topics. He also thanked all of the presenters for the first EDVMVAC meeting.

**Attachment A**

## EDMVAC Members in Attendance

Gerald LeBlanc, Chair
Thomas Osimitz, Co-Chair
Mildred Christian
Robert Combes
Rodger Curren
Peter DeFur
Anne Fairbrother
Paul Foster
David Hattan
Susan Jobling
William Kelce
Sean Kennedy
Steven Levine
Edward Orlando
James Owens
James Stevens
William Stokes
Glen Van Der Kraak